

## **STIC Biotechnology Systems Branch**

### **RAW SEQUENCE LISTING**

### **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/538,041  
Source: 1 FWP  
Date Processed by STIC: 2/6/06

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 4.4.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

**<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):  
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06

## Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: <u>10/538,041</u>
<b>ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE</b>		
1 _____ Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor <b>after</b> creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 _____ Invalid Line Length	The rules require that a line <b>not exceed</b> 72 characters in length. This includes white spaces.	
3 _____ Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. Do <b>not</b> use tab codes between numbers; use <b>space characters</b> , instead.	
4 _____ Non-ASCII	The submitted file was <b>not</b> saved in ASCII(DOS) text, as <b>required</b> by the Sequence Rules. Please <b>ensure your subsequent submission is saved in ASCII text.</b>	
5 <u>J</u> _____ Variable Length	Sequence(s) <u>123 (maybe more)</u> contain n's or Xaa's representing more than one residue. <b>Per Sequence Rules, each n or Xaa can only represent a single residue.</b> Please present the <b>maximum</b> number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 _____ PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. <b>This applies to the mandatory &lt;220&gt;-&lt;223&gt; sections for Artificial or Unknown sequences.</b>	
7 _____ Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for <b>each</b> skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to <b>include</b> the skipped sequences.	
8 _____ Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If <b>intentional</b> , please insert the following lines for <b>each</b> skipped sequence. <210> sequence id number <400> sequence id number 000	
9 _____ Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is <b>MANDATORY</b> if n's or Xaa's are present. In <220> to <223> section, please explain location of <b>n</b> or <b>Xaa</b> , and which residue <b>n</b> or <b>Xaa</b> represents.	
10 _____ Invalid <213> Response	Per 1.823 of Sequence Rules, the only <b>valid</b> <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is <b>required</b> when <213> response is Unknown or is Artificial Sequence	
11 _____ Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is <b>MANDATORY</b> if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 _____ PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 _____ Misuse of n/Xaa	<b>"n"</b> can <b>only</b> represent a single <u>nucleotide</u> ; <b>"Xaa"</b> can <b>only</b> represent a single <u>amino acid</u>	



IFWP

## RAW SEQUENCE LISTING

DATE: 02/06/2006

PATENT APPLICATION: US/10/538,041

TIME: 14:24:06

Input Set : A:\Sequence Listing for USSN 10-538041.txt

Output Set: N:\CRF4\02012006\J538041.raw

4 <110> APPLICANT: TOOLGEN, Inc.  
 6 <120> TITLE OF INVENTION: Regulatory Zinc Finger Proteins  
 8 <130> FILE REFERENCE: Q88285  
 10 <140> CURRENT APPLICATION NUMBER: US 10/538,041  
 11 <141> CURRENT FILING DATE: 2005-06-08  
 13 <150> PRIOR APPLICATION NUMBER: US 60/431,892  
 14 <151> PRIOR FILING DATE: 2002-12-09  
 16 <160> NUMBER OF SEQ ID NOS: 129  
 18 <170> SOFTWARE: PatentIn version 3.2  
 20 <210> SEQ ID NO: 1  
 21 <211> LENGTH: 23  
 22 <212> TYPE: PRT  
 23 <213> ORGANISM: Homo sapiens  
 25 <400> SEQUENCE: 1  
 27 Tyr Lys Cys Lys Gln Cys Gly Lys Ala Phe Gly Cys Pro Ser Asn Leu  
 28 1 5 10 15  
 30 Arg Arg His Gly Arg Thr His  
 31 20  
 34 <210> SEQ ID NO: 2  
 35 <211> LENGTH: 23  
 36 <212> TYPE: PRT  
 37 <213> ORGANISM: Homo sapiens  
 39 <400> SEQUENCE: 2  
 41 Tyr Ser Cys Gly Ile Cys Gly Lys Ser Phe Ser Asp Ser Ser Ala Lys  
 42 1 5 10 15  
 44 Arg Arg His Cys Ile Leu His  
 45 20  
 48 <210> SEQ ID NO: 3  
 49 <211> LENGTH: 23  
 50 <212> TYPE: PRT  
 51 <213> ORGANISM: Homo sapiens  
 53 <400> SEQUENCE: 3  
 55 Tyr Thr Cys Ser Asp Cys Gly Lys Ala Phe Arg Asp Lys Ser Cys Leu  
 56 1 5 10 15  
 58 Asn Arg His Arg Arg Thr His  
 59 20  
 62 <210> SEQ ID NO: 4  
 63 <211> LENGTH: 23  
 64 <212> TYPE: PRT  
 65 <213> ORGANISM: Homo sapiens  
 67 <400> SEQUENCE: 4  
 69 Tyr Lys Cys Gly Gln Cys Gly Lys Phe Tyr Ser Gln Val Ser His Leu  
 70 1 5 10 15

pp 6-8  
 Does Not Comply  
 Corrected Diskette Needed

## RAW SEQUENCE LISTING

DATE: 02/06/2006

PATENT APPLICATION: US/10/538,041

TIME: 14:24:06

Input Set : A:\Sequence Listing for USSN 10-538041.txt

Output Set: N:\CRF4\02012006\J538041.raw

72 Thr Arg His Gln Lys Ile His

73 20

76 &lt;210&gt; SEQ ID NO: 5

77 &lt;211&gt; LENGTH: 23

78 &lt;212&gt; TYPE: PRT

79 &lt;213&gt; ORGANISM: Homo sapiens

81 &lt;400&gt; SEQUENCE: 5

83 Tyr Lys Cys Glu Cys Gly Lys Ala Phe Arg Gln Ser Ser His Leu

84 1 5 10 15

86 Thr Thr His Lys Ile Ile His

87 20

90 &lt;210&gt; SEQ ID NO: 6

91 &lt;211&gt; LENGTH: 23

92 &lt;212&gt; TYPE: PRT

93 &lt;213&gt; ORGANISM: Homo sapiens

95 &lt;400&gt; SEQUENCE: 6

97 Tyr Glu Cys Glu Lys Cys Gly Lys Ala Phe Asn Gln Ser Ser Asn Leu

98 1 5 10 15

100 Thr Arg His Lys Lys Ser His

101 20

104 &lt;210&gt; SEQ ID NO: 7

105 &lt;211&gt; LENGTH: 23

106 &lt;212&gt; TYPE: PRT

107 &lt;213&gt; ORGANISM: Homo sapiens

109 &lt;400&gt; SEQUENCE: 7

111 Tyr Val Cys Ser Lys Cys Gly Lys Ala Phe Thr Gln Ser Ser Asn Leu

112 1 5 10 15

114 Thr Val His Gln Lys Ile His

115 20

118 &lt;210&gt; SEQ ID NO: 8

119 &lt;211&gt; LENGTH: 23

120 &lt;212&gt; TYPE: PRT

121 &lt;213&gt; ORGANISM: Homo sapiens

123 &lt;400&gt; SEQUENCE: 8

125 Tyr Lys Cys Pro Asp Cys Gly Lys Ser Phe Ser Gln Ser Ser Ser Leu

126 1 5 10 15

128 Ile Arg His Gln Arg Thr His

129 20

132 &lt;210&gt; SEQ ID NO: 9

133 &lt;211&gt; LENGTH: 25

134 &lt;212&gt; TYPE: PRT

135 &lt;213&gt; ORGANISM: Homo sapiens

137 &lt;400&gt; SEQUENCE: 9

139 Tyr Val Cys Asp Val Glu Gly Cys Thr Trp Lys Phe Ala Arg Ser Asp

140 1 5 10 15

142 Glu Leu Asn Arg His Lys Lys Arg His

143 20 25

146 &lt;210&gt; SEQ ID NO: 10

147 &lt;211&gt; LENGTH: 23

## RAW SEQUENCE LISTING

DATE: 02/06/2006

PATENT APPLICATION: US/10/538,041

TIME: 14:24:06

Input Set : A:\Sequence Listing for USSN 10-538041.txt

Output Set: N:\CRF4\02012006\J538041.raw

```

148 <212> TYPE: PRT
149 <213> ORGANISM: Homo sapiens
151 <400> SEQUENCE: 10
154 Phe Gln Cys Lys Thr Cys Gln Arg Lys Phe Ser Arg Ser Asp His Leu
155 1 5 10 15
157 Lys Thr His Thr Arg Thr His
158 20
161 <210> SEQ ID NO: 11
162 <211> LENGTH: 23
163 <212> TYPE: PRT
164 <213> ORGANISM: Homo sapiens
166 <400> SEQUENCE: 11
168 Tyr Lys Cys Met Glu Cys Gly Lys Ala Phe Asn Arg Arg Ser His Leu
169 1 5 10 15
171 Thr Arg His Gln Arg Ile His
172 20
175 <210> SEQ ID NO: 12
176 <211> LENGTH: 23
177 <212> TYPE: PRT
178 <213> ORGANISM: Homo sapiens
180 <400> SEQUENCE: 12
182 Tyr Ile Cys Arg Lys Cys Gly Arg Gly Phe Ser Arg Lys Ser Asn Leu
183 1 5 10 15
185 Ile Arg His Gln Arg Thr His
186 20
189 <210> SEQ ID NO: 13
190 <211> LENGTH: 23
191 <212> TYPE: PRT
192 <213> ORGANISM: Homo sapiens
194 <400> SEQUENCE: 13
196 Tyr Glu Cys Asp His Cys Gly Lys Ala Phe Ser Val Ser Ser Asn Leu
197 1 5 10 15
199 Asn Val His Arg Arg Ile His
200 20
204 <210> SEQ ID NO: 14
205 <211> LENGTH: 23
206 <212> TYPE: PRT
207 <213> ORGANISM: Homo sapiens
209 <400> SEQUENCE: 14
211 Tyr Thr Cys Lys Gln Cys Gly Lys Ala Phe Ser Val Ser Ser Ser Leu
212 1 5 10 15
214 Arg Arg His Glu Thr Thr His
215 20
218 <210> SEQ ID NO: 15
219 <211> LENGTH: 23
220 <212> TYPE: PRT
221 <213> ORGANISM: Homo sapiens
223 <400> SEQUENCE: 15
225 Tyr Glu Cys Asn Tyr Cys Gly Lys Thr Phe Ser Val Ser Ser Thr Leu

```

## RAW SEQUENCE LISTING

DATE: 02/06/2006

PATENT APPLICATION: US/10/538,041

TIME: 14:24:06

Input Set : A:\Sequence Listing for USSN 10-538041.txt

Output Set: N:\CRF4\02012006\J538041.raw

```

226 1           5           10           15
228 Ile Arg His Gln Arg Ile His
229           20
232 <210> SEQ ID NO: 16
233 <211> LENGTH: 23
234 <212> TYPE: PRT
235 <213> ORGANISM: Homo sapiens
237 <400> SEQUENCE: 16
239 Tyr Arg Cys Glu Glu Cys Gly Lys Ala Phe Arg Trp Pro Ser Asn Leu
240 1           5           10           15
242 Thr Arg His Lys Arg Ile His
243           20
246 <210> SEQ ID NO: 17
247 <211> LENGTH: 23
248 <212> TYPE: PRT
249 <213> ORGANISM: Homo sapiens
251 <400> SEQUENCE: 17
254 Tyr Glu Cys Asp His Cys Gly Lys Ser Phe Ser Gln Ser Ser His Leu
255 1           5           10           15
257 Asn Val His Lys Arg Thr His
258           20
261 <210> SEQ ID NO: 18
262 <211> LENGTH: 23
263 <212> TYPE: PRT
264 <213> ORGANISM: Homo sapiens
266 <400> SEQUENCE: 18
268 Phe Leu Cys Gln Tyr Cys Ala Gln Arg Phe Gly Arg Lys Asp His Leu
269 1           5           10           15
271 Thr Arg His Met Lys Lys Ser
272           20
275 <210> SEQ ID NO: 19
276 <211> LENGTH: 24
277 <212> TYPE: PRT
278 <213> ORGANISM: Artificial
280 <220> FEATURE:
281 <223> OTHER INFORMATION: Artificial zinc finger domain
283 <400> SEQUENCE: 19
285 Tyr Arg Cys Lys Tyr Cys Asp Arg Ser Phe Ser Asp Ser Ser Asn Leu
286 1           5           10           15
288 Gln Arg His Val Arg Asn Ile His
289           20
292 <210> SEQ ID NO: 20
293 <211> LENGTH: 83
294 <212> TYPE: PRT
295 <213> ORGANISM: Artificial
297 <220> FEATURE:
298 <223> OTHER INFORMATION: artificial zinc finger protein
300 <400> SEQUENCE: 20
302 Tyr Lys Cys Gly Gln Cys Gly Lys Phe Tyr Ser Gln Val Ser His Leu

```

## RAW SEQUENCE LISTING

DATE: 02/06/2006

PATENT APPLICATION: US/10/538,041

TIME: 14:24:06

Input Set : A:\Sequence Listing for USSN 10-538041.txt

Output Set: N:\CRF4\02012006\J538041.raw

```

303 1           5           10           15
304 Thr Arg His Gln Lys Ile His Thr Gly Glu Lys Pro Phe Gln Cys Lys
305           20           25           30
307 Thr Cys Gln Arg Lys Phe Ser Arg Ser Asp His Leu Lys Thr His Thr
308           35           40           45
310 Arg Thr His Thr Gly Glu Lys Pro Tyr Ile Cys Arg Lys Cys Gly Arg
311           50           55           60
313 Gly Phe Ser Arg Lys Ser Asn Leu Ile Arg His Gln Arg Thr His Thr
314 65           70           75           80
316 Gly Glu Lys
320 <210> SEQ ID NO: 21
321 <211> LENGTH: 83
322 <212> TYPE: PRT
323 <213> ORGANISM: Artificial
325 <220> FEATURE:
326 <223> OTHER INFORMATION: artificial zinc finger protein
328 <400> SEQUENCE: 21
330 Tyr Lys Cys Glu Glu Cys Gly Lys Ala Phe Arg Gln Ser Ser His Leu
331 1           5           10           15
333 Thr Thr His Lys Ile Ile His Thr Gly Glu Lys Pro Tyr Lys Cys Met
334           20           25           30
336 Glu Cys Gly Lys Ala Phe Asn Arg Arg Ser His Leu Thr Arg His Gln
337           35           40           45
339 Arg Ile His Thr Gly Glu Lys Pro Phe Gln Cys Lys Thr Cys Gln Arg
340           50           55           60
342 Lys Phe Ser Arg Ser Asp His Leu Lys Thr His Thr Arg Thr His Thr
343 65           70           75           80
345 Gly Glu Lys
349 <210> SEQ ID NO: 22
350 <211> LENGTH: 83
351 <212> TYPE: PRT
352 <213> ORGANISM: Artificial
354 <220> FEATURE:
355 <223> OTHER INFORMATION: artificial zinc finger protein
357 <400> SEQUENCE: 22
359 Tyr Lys Cys Gly Gln Cys Gly Lys Phe Tyr Ser Gln Val Ser His Leu
360 1           5           10           15
362 Thr Arg His Gln Lys Ile His Thr Gly Glu Lys Pro Phe Gln Cys Lys
363           20           25           30
365 Thr Cys Gln Arg Lys Phe Ser Arg Ser Asp His Leu Lys Thr His Thr
366           35           40           45
368 Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Met Glu Cys Gly Lys
369           50           55           60
371 Ala Phe Asn Arg Arg Ser His Leu Thr Arg His Gln Arg Ile His Thr
372 65           70           75           80
374 Gly Glu Lys
378 <210> SEQ ID NO: 23
379 <211> LENGTH: 83
380 <212> TYPE: PRT

```

10/538,041 6

<210> 123  
<211> 21  
<212> PRT  
<213> Artificial  
  
<220>  
<223> zinc finger consensus

<220>  
<221> MISC\_FEATURE  
<222> (2)..(2)  
<223> any amino acid

<220>  
<221> MISC\_FEATURE  
<222> (3)..(3)  
<223> between 1 and 4 amino acids of any amino acid

↓  
variable length not permitted - see item 5  
on Eran  
summary sheet



from sequence 123

10/538,041 9

<220>

<221> MISC\_FEATURE

<222> (20)..(20)

<223> between one and three residues of any amino acid

<400> 123

same error

Please correct this type of error  
in subsequent sequences

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/538,041

DATE: 02/06/2006  
TIME: 14:24:07

Input Set : A:\Sequence Listing for USSN 10-538041.txt  
Output Set: N:\CRF4\02012006\J538041.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:123; Xaa Pos. 2,3,5,6,7,8,9,10,11,12,13,14,15,16,18,19,20

Seq#:124; Xaa Pos. 2,3,5,6,7,8,9,11,14,15,18,19,20

Seq#:125; Xaa Pos. 3,4,6

Invalid <213> Responses

Use of "Artificial" only as "<213> Organism" response is incomplete,  
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42

Seq#:43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66

Seq#:67,68,69,70,71,123,124,125,126,127,128

**VERIFICATION SUMMARY**

DATE: 02/06/2006

PATENT APPLICATION: US/10/538,041

TIME: 14:24:07

Input Set : A:\Sequence Listing for USSN 10-538041.txt

Output Set: N:\CRF4\02012006\J538041.raw

L:3279 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123 after pos.:0  
M:341 Repeated in SeqNo=123  
L:3304 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3351 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:124 after pos.:0  
M:341 Repeated in SeqNo=124  
L:3384 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:125 after pos.:0